

IBM SPSS Bootstrapping 30

IBM

Nota

Prima di utilizzare queste informazioni e il prodotto che supportano, leggere le informazioni in “Informazioni particolari” a pagina 7.

Informazioni sul prodotto

Questa edizione si applica alla versione 30, release 0, modifica 0 di IBM® SPSS Statistiche e a tutte le release e modifiche successive se non diversamente indicato in nuove edizioni.

© **Copyright International Business Machines Corporation .**

Indice

| | |
|---|-----------|
| Capitolo 1. Bootstrap..... | 1 |
| Introduzione al bootstrap..... | 1 |
| Bootstrap..... | 1 |
| Procedure che supportano bootstrap..... | 2 |
| Funzioni aggiuntive di comando BOOTSTRAP..... | 5 |
| Informazioni particolari..... | 7 |
| Marchi..... | 8 |
| Indice analitico..... | 11 |

Capitolo 1. Bootstrap

Le seguenti funzioni di bootstrap sono incluse nella Base Edition.

Introduzione al bootstrap

Quando si raccolgono i dati, si è spesso interessati alle proprietà della popolazione da cui si è preso il campione. Si fanno inferenze su questi parametri di popolazione con stime calcolate dal campione. Ad esempio, se il dataset *Employee data.sav* incluso con il prodotto è un campione casuale proveniente da una popolazione più ampia di dipendenti, la media del campione di \$34,419.57 per *Stipendio corrente* è una stima dello stipendio corrente medio per la popolazione di dipendenti. Inoltre, questa stima ha un errore standard di \$784.311 per un campione di dimensione 474 e quindi un intervallo di confidenza del 95% per lo stipendio medio corrente nella popolazione di dipendenti è di \$32, da \$78.40 a \$35,960.73. Ma quanto sono affidabili questi estimatori? Per alcune popolazioni "conosciute" e parametri ben comportamentali, sappiamo abbastanza bene le proprietà delle stime del campione, e possono essere fiduciosi in questi risultati. Bootstrap cerca di scoprire maggiori informazioni sulle proprietà degli estimatori per le popolazioni "sconosciute" e i parametri mal comportati.

Come funziona bootstrap

Al suo modo più semplice, per un dataset con una dimensione di esempio di N , si prendono i campioni di bootstrap B di dimensione N con sostituzione dal dataset originale e compatta l'estimatore per ciascuno di questi campioni di bootstrap B . Queste stime di bootstrap B sono un esempio di dimensione B da cui si possono fare inferenze sullo stimatore. Ad esempio, se si prelevano 1.000 campioni di bootstrap dal set di dati *Employee data.sav*, l'errore standard stimato bootstrap di \$776.91 per la media del campione per *Stipendio corrente* è un'alternativa alla stima di \$784.311.

Inoltre, il bootstrap fornisce un errore standard e un intervallo di confidenza per la mediana, per cui le stime parametriche non sono disponibili.

Supporto per la bootstrap nel prodotto

Bootstrap è incorporato come sottofinestra di dialogo nelle procedure che supportano lo bootstrap. Consultare ["Procedure che supportano bootstrap"](#) a pagina 2 per informazioni su quali procedure supportano bootstrap.

Quando viene richiesto il bootstrap nelle finestre di dialogo, viene incollato un nuovo e separato comando `BOOTSTRAP` oltre alla consueta sintassi generata dalla finestra di dialogo. Il comando `BOOTSTRAP` crea i campioni di bootstrap in base alle tue specifiche. Internamente il prodotto tratta questi campioni di bootstrap come splits, anche se non vengono mostrati esplicitamente nell'editor dei dati. Ciò significa che, internamente, ci sono effettivamente i casi $B*N$, quindi il contatore del caso nella barra di stato conterà da 1 a $B*N$ quando si elaborano i dati durante la bootstrap. Il sistema di gestione dell'output (OMS) viene utilizzato per raccogliere i risultati di esecuzione dell'analisi su ogni "split bootstrap". Questi risultati sono pooled e i risultati bootstrap pooled visualizzati nel Viewer con il resto della consueta uscita generata dalla procedura. In certi casi, si può vedere un riferimento a "bootstrap split 0"; questo è il dataset originale.

Nota: I grafici non vengono prodotti nell'output quando il bootstrap è abilitato.

Bootstrap

Bootstrap è un metodo per derivare stime robuste di errori standard e intervalli di confidenza per stime come la media, la mediana, la proporzione, il rapporto di probabilità, il coefficiente di correlazione o il coefficiente di regressione. Può essere utilizzato anche per creare dei test sull'ipotesi. Il bootstrap è più utile come alternativa alle stime parametriche quando le ipotesi di quei metodi sono in dubbio (come nel caso dei modelli di regressione con residui eteroscedastici si adattano a piccoli campioni), oppure dove

l'inferenza parametrica è impossibile o richiede formule molto complicate per il calcolo di errori standard (come nel caso di intervalli di confidenza di calcolo per la mediana, quartili e altri percentili).

Esempi

Un'azienda di telecomunicazioni perde circa il 27% dei suoi clienti per indursi ogni mese. Al fine di concentrare correttamente gli sforzi di riduzione della chiesta, la gestione vuole sapere se questa percentuale varia tra i gruppi di clienti predefiniti. Utilizzando bootstrap, è possibile stabilire se una singola velocità di churn descrive adeguatamente i quattro principali tipi di clienti.

In una revisione dei record dei dipendenti, la gestione è interessata alla precedente esperienza lavorativa dei dipendenti. L'esperienza lavorativa è corretta, il che rende la media una stima meno desiderabile della precedente esperienza lavorativa "tipica" tra i dipendenti rispetto alla mediana. Tuttavia, gli intervalli di confidenza parametrica non sono disponibili per la mediana nel prodotto.

La gestione è interessata anche a determinare quali fattori sono associati agli aumenti di stipendio dei dipendenti, adattando un modello lineare alla differenza tra gli stipendi attuali e quelli di partenza. Quando si bootstrap un modello lineare, è possibile utilizzare metodi di ri - campionamento speciali (bootstrap residui e selvatici) per ottenere risultati più accurati.

Molte procedure supportano il campionamento bootstrap e il pooling dei risultati dall'analisi dei campioni di bootstrap. I controlli per la specifica delle analisi di bootstrap sono integrati direttamente come un comune sottodialogo nelle procedure che supportano la bootstrap. Le impostazioni sulla finestra di dialogo bootstrap persistono tra le procedure in modo che se si esegue un'analisi delle frequenze con bootstrap tramite le finestre di dialogo, il bootstrap verrà attivato per impostazione predefinita per altre procedure che lo supportano.

Ottenimento di un'analisi bootstrap

1. Dai menu scegliere una procedura che supporta bootstrap e clicca su **Bootstrap**.
2. Selezionare **Perform bootstrap**.

Facoltativamente, è possibile controllare le seguenti opzioni:

Numero di campioni. Per gli intervalli percentili e BCa prodotti, si consiglia di utilizzare almeno 1000 campioni di bootstrap. Specificare un intero positivo.

Imposta seme per Mersenne Twister. Impostando un seme è possibile replicare le analisi. L'utilizzo di questo controllo è simile all'impostazione di Mersenne Twister come generatore attivo e alla specifica di un punto iniziale fisso nella finestra di dialogo Generatori di numeri casuali, con la differenza importante che impostando il seme in questa finestra di dialogo preserverà lo stato corrente del generatore di numeri casuali e ripristinerà tale stato una volta completata l'analisi.

Intervalli di confidenza. Specificare un livello di confidenza maggiore di 50 e minore di 100. Gli intervalli percentili utilizzano semplicemente i valori di bootstrap ordinati corrispondenti all'intervallo di confidenza percentili. Ad esempio, un intervallo di confidenza di 95% percentile utilizza i percentili 2.5th e 97.5th dei valori bootstrap come limiti inferiori e superiori dell'intervallo (interpolando i valori di bootstrap se necessario). Gli intervalli di Bias corretti e accelerati (BCa) sono intervalli regolati più accurati al costo di richiedere più tempo al compimento.

Campionamento. Il metodo **Semplice** è il caso di ri - campionamento con sostituzione dal dataset originale. Il metodo **Stratified** è il caso ri - sampling con sostituzione dal dataset originale, *entro* gli strati definiti dalla classificazione incrociata delle variabili strati. Il campionamento di bootstrap stratificato può essere utile quando le unità all'interno di strati sono relativamente omogenee mentre le unità su strati sono molto diverse.

Nota: I grafici non vengono prodotti nell'output quando il bootstrap è abilitato.

Procedure che supportano bootstrap

Le seguenti procedure supportano bootstrap.

- Bootstrap non funziona con datasets moltiplicati. Se c'è una variabile *Imputation_* nel dataset, la finestra di dialogo Bootstrap è disabilitata.
- Bootstrap non funziona se ci sono valori di peso non interi.
- Bootstrapping utilizza la cancellazione listwise per determinare la base del caso; cioè i casi con valori mancanti su una qualsiasi delle variabili di analisi vengono eliminati dall'analisi, quindi quando è in atto la bootstrap, la cancellazione listwise è in vigore anche se la procedura di analisi specifica un'altra forma di gestione del valore mancante.

Nota: I grafici generalmente non vengono forniti quando il bootstrap è in vigore per un'analisi.

Edizione base statistiche

Frequenze

Sono supportate le seguenti funzioni:

- La tabella Statistiche supporta le stime di bootstrap per la media, deviazione standard, varianza, mediana, skewness, kurtosi e percentili.
- La Tabella Frequenze supporta le stime di bootstrap per la percentuale.

Descrittive

Sono supportate le seguenti funzioni:

- La Tabella Statistica Descrittiva supporta le stime di bootstrap per la media, la deviazione standard, la varianza, la skewness e la kurtosi.

Esplora

Sono supportate le seguenti funzioni:

- La Tabella Descrittiva supporta le stime di bootstrap per la media, 5% Trimmed Mean, deviazione standard, varianza, mediana, skewness, kurtosi e range interquartile.
- La M-Estimators table supporta le stime di bootstrap per il M-Estimator di Huber, il Biweight di Tukey, il M-Estimator di Hampel e il Wave di Andrew.
- La tabella Percentili supporta le stime di bootstrap per i percentili.

Tabelle di contingenza

Sono supportate le seguenti funzioni:

- La tabella Misure direzionali supporta le stime di bootstrap per Lambda, Goodman e Kruskal Tau, Incertezza Coefficiente e Somers ' d.
- La tabella Misure Symmetric supporta le stime di bootstrap per Phi, Cramer's V, Contingency Coefficiente, Kendall's tau - b, Kendall's tau - c, Gamma, Spearman Correlation e Pearson's R.
- Il Risk Preventivo table supporta le stime di bootstrap per il rapporto di odds.
- Il Mantel-Haenszel Common Odds Ratio table supporta le stime di bootstrap e i test di significatività per ln (Stima).

Medie

Sono supportate le seguenti funzioni:

- La tabella Report supporta le stime di bootstrap per la media, mediana, mediana raggruppata, deviazione standard, varianza, kurtosi, skewness, media armonica e media geometrica.

Test T a campione singolo

Sono supportate le seguenti funzioni:

- La tabella Statistiche supporta le stime di bootstrap per la deviazione media e standard.
- La tabella Test supporta le stime di bootstrap e i test di significatività per la differenza media.

Test T a campioni indipendenti

Sono supportate le seguenti funzioni:

- La tabella Statistiche del gruppo supporta le stime di bootstrap per la deviazione media e standard.

- La tabella Test supporta le stime di bootstrap e i test di significatività per la differenza media.

Test T a campioni accoppiati

Sono supportate le seguenti funzioni:

- La tabella Statistiche supporta le stime di bootstrap per la deviazione media e standard.
- La tabella Correlazioni supporta le stime di bootstrap per le correlazioni.
- La tabella Test supporta le stime di bootstrap per la media.

ANOVA a una via

Sono supportate le seguenti funzioni:

- La Tabella Statistica Descrittiva supporta le stime di bootstrap per la deviazione media e standard.
- La Tabella Multiple Confronti supporta le stime di bootstrap per la differenza media.
- Il Tavolo Contrasto Test supporta le stime di bootstrap e i test di significazione per il valore del contrasto.

GLM univariato

Sono supportate le seguenti funzioni:

- La Tabella Statistica Descrittiva supporta le stime di bootstrap per la deviazione Mean e standard.
- La tabella delle stime dei parametri supporta le stime di bootstrap e i test di significatività per il coefficiente, B.
- La Tabella Dei Risultati di contrasto supporta le stime di bootstrap e i test di significazione per la differenza.
- Il Significato marginale stimato: tabella di stime supporta le stime di bootstrap per la media.
- Il Significato Marginale Stimato: Pairwise Confronti table supporta le stime di bootstrap per la differenza media.
- I Test Post Hoc: Molteplici Confronti da tavolo supporta le stime di bootstrap per la Mean Difference.

Correlazioni di bivariate

Sono supportate le seguenti funzioni:

- La Tabella Statistica Descrittiva supporta le stime di bootstrap per la deviazione media e standard.
- La tabella Correlazioni supporta le stime di bootstrap e i test di significatività per le correlazioni.

Nota: Se le correlazioni nonparametriche (Kendall's tau - b o Spearman) vengono richieste in aggiunta alle correlazioni di Pearson, i comandi di dialogo pastes CORRELATIONS e NONPAR CORR con un comando B00TSTRAP separato per ciascuno. Gli stessi campioni di bootstrap saranno utilizzati per calcolare tutte le correlazioni.

Prima del pooling, la trasformata di Fisher Z viene applicata alle correlazioni. Dopo il pooling, viene applicata la trasformazione inversa Z .

Correlazioni parziali

Sono supportate le seguenti funzioni:

- La Tabella Statistica Descrittiva supporta le stime di bootstrap per la deviazione media e standard.
- La tabella Correlazioni supporta le stime di bootstrap per le correlazioni.

Regressione lineare

Sono supportate le seguenti funzioni:

- La Tabella Statistica Descrittiva supporta le stime di bootstrap per la deviazione media e standard.
- La tabella Correlazioni supporta le stime di bootstrap per le correlazioni.
- La tabella Riepilogo modelli supporta le stime di bootstrap per Durbin -Watson.
- La tabella dei coefficienti supporta le stime di bootstrap e i test di significatività per il coefficiente, B.
- La Tabella Correlazione Coefficienti supporta le stime di bootstrap per le correlazioni.
- La Tabella Residui Statistiche supporta le stime di bootstrap per la deviazione media e standard.

Regressione ordinale

Sono supportate le seguenti funzioni:

- La tabella delle stime dei parametri supporta le stime di bootstrap e i test di significatività per il coefficiente, B.

Analisi discriminante

Sono supportate le seguenti funzioni:

- Il Tavolo Standardized Discriminant Function Coefficients table supporta le stime di bootstrap per i coefficienti standardizzati.
- Il Canonical Discriminant Function Coefficients table supporta le stime di bootstrap per i coefficienti non standardizzati.
- La tabella Coefficienti di classificazione delle funzioni supporta le stime di bootstrap per i coefficienti.

SPSS Statistiche Edizione Premium e Opzione statistiche avanzate

GLM Multivariato

Sono supportate le seguenti funzioni:

- La tabella delle stime dei parametri supporta le stime di bootstrap e i test di significatività per il coefficiente, B.

Modelli misti lineari

Le seguenti funzioni sono supportate a meno che non sia specificato un comando RIPETUTO:

- Le stime della tabella Fixed Effects supporta le stime di bootstrap e i test di significatività per la stima.
- Le stime della tabella dei parametri di Covariance supporta le stime di bootstrap e i test di significatività per la stima.

Modelli lineari generalizzati

Sono supportate le seguenti funzioni:

- La tabella delle stime dei parametri supporta le stime di bootstrap e i test di significatività per il coefficiente, B.

Regressione di Cox

Sono supportate le seguenti funzioni:

- Le Variabili nella tabella Equitazione supporta le stime di bootstrap e i test di significazione per il coefficiente, B.

SPSS Statistiche Standard Edition e Opzione di regressione

Regressione logistica binaria

Sono supportate le seguenti funzioni:

- Le Variabili nella tabella Equitazione supporta le stime di bootstrap e i test di significazione per il coefficiente, B.

Regressione logistica multinomiale

Sono supportate le seguenti funzioni:

- La tabella delle stime dei parametri supporta le stime di bootstrap e i test di significatività per il coefficiente, B.

Funzioni aggiuntive di comando BOOTSTRAP

Il linguaggio della sintassi dei comandi consente inoltre di:

- Eseguire il campionamento del bootstrap residuo e wild (SAMPLING secondario)

Vedere *Command Syntax Reference* per informazioni dettagliate sulla sintassi.

Informazioni particolari

Queste informazioni sono state sviluppate per prodotti e servizi offerti negli Stati Uniti. Questo materiale potrebbe essere disponibile da IBM in altre lingue. Tuttavia, può essere necessario possedere una copia del prodotto o versione del prodotto in tale lingua per potervi accedere.

IBM può non offrire i prodotti, i servizi o le funzioni presentati in questo documento in altri paesi. Consultare il proprio rappresentante locale IBM per informazioni sui prodotti ed i servizi attualmente disponibili nella propria zona. Qualsiasi riferimento ad un prodotto, programma o servizio IBM non implica o intende dichiarare che solo quel prodotto, programma o servizio IBM può essere utilizzato. Qualsiasi prodotto funzionalmente equivalente al prodotto, programma o servizio che non violi alcun diritto di proprietà intellettuale IBM può essere utilizzato. È comunque responsabilità dell'utente valutare e verificare la possibilità di utilizzare altri prodotti, programmi o servizi non IBM.

IBM può avere applicazioni di brevetti o brevetti in corso relativi all'argomento descritto in questo documento. La fornitura del presente documento non concede alcuna licenza a tali brevetti. È possibile inviare per iscritto richieste di licenze a:

IBM Director of Licensing
IBM Corporation
North Castle Drive, MD-NC119
Armonk, NY 10504-1785
US

Per richieste di licenze relative ad informazioni double-byte (DBCS), contattare il Dipartimento di Proprietà Intellettuale IBM nel proprio paese o inviare richieste per iscritto a:

Intellectual Property Licensing
Legal and Intellectual Property Law
IBM Japan Ltd.
19-21, Nihonbashi-Hakozakicho, Chuo-ku
Tokyo 103-8510, Japan

IBM (INTERNATIONAL BUSINESS MACHINES CORPORATION) FORNISCE LA PRESENTE PUBBLICAZIONE "NELLO STATO IN CUI SI TROVA" SENZA GARANZIE DI ALCUN TIPO, ESPRESSE O IMPLICITE, IVI INCLUSE, A TITOLO DI ESEMPIO, GARANZIE IMPLICITE DI NON VIOLAZIONE, DI COMMERCIALIZZABILITÀ E DI IDONEITÀ PER UNO SCOPO PARTICOLARE. Alcune giurisdizioni non consentono la rinuncia ad alcune garanzie espresse o implicite in determinate transazioni, pertanto, la presente dichiarazione può non essere applicabile.

Queste informazioni potrebbero contenere imprecisioni tecniche o errori tipografici. Le modifiche alle presenti informazioni vengono effettuate periodicamente; tali modifiche saranno incorporate nelle nuove pubblicazioni della pubblicazione. IBM si riserva il diritto di apportare miglioramenti e/o modifiche al prodotto o al programma descritto nel manuale in qualsiasi momento e senza preavviso.

I riferimenti in queste informazioni a siti Web non IBM vengono forniti solo per comodità e non implicano in alcun modo l'approvazione di tali siti web. I materiali presenti in tali siti web non sono parte dei materiali per questo prodotto IBM e l'utilizzo di tali siti è a proprio rischio.

IBM può utilizzare o distribuire qualsiasi informazione fornita in qualsiasi modo ritenga appropriato senza incorrere in alcun obbligo verso l'utente.

Coloro che detengano la licenza su questo programma e desiderano avere informazioni su di esso allo scopo di consentire: (i) uno scambio di informazioni tra programmi indipendenti ed altri (compreso questo) e (ii) l'utilizzo reciproco di tali informazioni, dovrebbe rivolgersi a:

IBM Director of Licensing
IBM Corporation
North Castle Drive, MD-NC119

Armonk, NY 10504-1785

US

Tali informazioni possono essere disponibili, in base ad appropriate clausole e condizioni, includendo in alcuni casi, il pagamento di una tassa.

Il programma concesso in licenza descritto nel presente documento e tutto il materiale concesso in licenza disponibile sono forniti da IBM in base alle clausole dell'Accordo per Clienti IBM (IBM Customer Agreement), dell'IBM IPLA (IBM International Program License Agreement) o qualsiasi altro accordo equivalente tra le parti.

Gli esempi client e dati delle prestazioni citati sono presentati esclusivamente a scopo illustrativo. Gli effettivi risultati delle prestazioni possono variare in base alle configurazioni e alle condizioni operative specifiche.

Le informazioni relative a prodotti non IBM sono ottenute dai fornitori di quei prodotti, dagli annunci pubblicati o da altre fonti disponibili al pubblico. IBM non ha testato quei prodotti e non può confermarne la precisione della prestazione, la compatibilità o qualsiasi altro reclamo relativo ai prodotti non IBM. Le domande sulle funzionalità dei prodotti non IBM devono essere indirizzate ai fornitori di tali prodotti.

Le dichiarazioni relative all'orientamento o alle intenzioni future di IBM sono soggette a modifica o a ritiro senza preavviso e rappresentano solo mete e obiettivi.

Questa pubblicazione contiene esempi di dati e prospetti utilizzati quotidianamente nelle operazioni aziendali. Pertanto, per maggiore completezza, gli esempi includono nomi di persone, società, marchi e prodotti. Tutti i nomi contenuti nel manuale sono fittizi e ogni riferimento a persone o aziende reali è puramente casuale.

LICENZA DI COPYRIGHT:

Queste informazioni contengono programmi applicativi di esempio in linguaggio sorgente, che illustrano tecniche di programmazione su varie piattaforme operative. È possibile copiare, modificare e distribuire questi programmi di esempio sotto qualsiasi forma senza alcun pagamento a IBM, allo scopo di sviluppare, utilizzare, commercializzare o distribuire i programmi applicativi in conformità alle API (application programming interface) a seconda della piattaforma operativa per cui i programmi di esempio sono stati scritti. Questi esempi non sono stati testati approfonditamente tenendo conto di tutte le condizioni possibili. IBM, quindi, non può garantire o sottintendere l'affidabilità, l'utilità o il funzionamento di questi programmi. I programmi di esempio sono forniti "NELLO STATO IN CUI SI TROVANO", senza alcun tipo di garanzia. IBM non intende essere responsabile per alcun danno derivante dall'uso dei programmi di esempio.

Ogni copia o qualsiasi parte di questi programmi di esempio o qualsiasi lavoro derivato, devono contenere le seguenti informazioni relative alle leggi sul diritto d'autore:

© Copyright IBM Corp. 2021. Le porzioni di questo codice derivano da IBM Corp. Programmi di esempio.

© Copyright IBM Corp. 1989 - 2021. Tutti i diritti riservati.

Marchi

IBM, il logo IBM e ibm.com sono marchi o marchi registrati di International Business Machines Corp., registrati in molte giurisdizioni in tutto il mondo. Altri nomi di prodotti e servizi possono essere marchi di IBM o di altre società. Un elenco corrente dei marchi IBM è disponibile sul web in "Copyright and trademark information" all'indirizzo www.ibm.com/legal/copytrade.shtml.

Adobe, il logo Adobe, PostScript e il logo PostScript sono marchi o marchi registrati di Adobe Systems Incorporated negli Stati Uniti e/o in altri paesi.

Intel, il logo Intel, Intel Inside, il logo Intel Inside, Intel Centrino, il logo Intel Centrino, Celeron, Intel Xeon, Intel SpeedStep, Itanium e Pentium sono marchi o marchi registrati di Intel Corporation o delle sue consociate negli Stati Uniti e in altri paesi.

Linux è un marchio registrato di Linus Torvalds negli Stati Uniti e/o negli altri paesi.

Microsoft, Windows, Windows NT e il logo Windows sono marchi di Microsoft Corporation negli Stati Uniti e/o negli altri paesi.

UNIX è un marchio della The Open Group negli Stati Uniti e/o negli altri paesi.

Java e tutti i marchi e i logo basati su Java sono marchi o marchi registrati di Oracle e/o associate.

Indice analitico

A

avvio

procedure supportate 2

